Review

Let $X_1, X_2, \cdots \sim^{\text{ind}} F$, where F has mean and variance

$$\mu = \int_{-\infty}^{\infty} x dF(x)$$

$$\sigma^2 = \int_{-\infty}^{\infty} (x - mu)^2 dF(x).$$

The LLN

$$\frac{X_1 + \cdots + X_n}{n} \to \mu.$$

The CLThm: Let $S_n = X_1 + \cdots + X_n$. Then

$$\lim_{n\to\infty} P\left[\frac{S_n - n\mu}{\sigma\sqrt{n}} \le z\right] = \Phi(z),$$

where Φ is the standard normal distribution function; that is,

$$S_n \approx \text{Normal}(n\mu, n\sigma^2).$$

for large n.

Today: An Application: Simulation.

Prologue

Let

$$A, B, C \sim^{ind} \text{Unif}[-L, L].$$

What is the probability that

$$Ax^2 + Bx + C = 0$$

has real roots (in x)? This is

$$P[4AC \le B^2] = \frac{1}{8L^3} \int\int\int_R dadbdc,$$

where

$$R = \{(a, b, c) : -L \le a, b, c \le L,\ 4ac \le b^2\}.$$

Notes:
- Can be computed
- With difficulty
- Does not depend on L.

The LLN For Events

If

$$E_1, E_2, \cdots \text{ are independent,}$$
$$P(E_i) = p,$$

then

$$\frac{\#\{k \le N : E_k \text{ occurs}\}}{N} \to p.$$

Notes:
- Qualitative
- CLThm

Applications:
- Prediction–e.g. gambling.
- Inference–today.
- Use to compute p.

Simulation

The LLN In Action

Outline

- Random Number Generators
- Simulation Experiments
- Examples
- Non-uniform Distributions

Modular Arithmetic

If m is any positive integer, then any integer n may be written

$$n = km + r,$$

where k and r are integers for which

$$0 \leq r < m.$$

Write

$$n = r \pmod{m}.$$

If

$$n_i = k_i m + r_i,$$

then $n_1 = n_2 \pmod{m}$ if $r_1 = r_2$.

Example: $10 = 1$ (mod3).

Linear Congruential Generators

Consider the recursion

$$X_{n+1} = (aX_n + c) \pmod{m},$$

where $a,\ c,\ m$, and X_0 are positive integers (parameters of the LCG). Here

$$a = \text{multiplier},$$
$$c = \text{increment},$$
$$m = \text{modulus},$$
$$X_0 = \text{seed}.$$

Example: With $X_0 = 1,\ a = 3,\ c = 1$, and $m = 8$,

$$X_1 = 3 \times 1 + 1 = 4 \text{ (mod8)}$$
$$X_2 = 3 \times 4 + 1 = 5 \text{ (mod8)}$$
$$X_3 = 3 \times 5 + 1 = 0 \text{ (mod8)}$$
$$X_4 = 3 \times 0 + 1 = 1 \text{ (mod8)}$$
$$\cdots,$$

Pseudo Random Number Generators

For *suitable* $a,\ c,\ m$, and X_0,

$$U_n = \frac{X_n}{m}$$

simulate independent standard uniform random variables Unif$[0, 1]$.

Example

$$a = 129,$$
$$c = 1,$$
$$m = 2^{35}.$$

Note: RAND or RND.

Example

Q: $A, B, C \sim^{ind}$ Unif$[-L, L]$, what is the probability that $Ax^2 + Bx + C = 0$ has real roots (in x)?

Ans: Generate

$$A_1, B_1, C_1, A_2, \cdots, B_N, C_N \sim^{ind} \text{Unif}[-1, 1].$$

Let

$$E_i = \{4A_iC_i \leq B_i^2\}.$$

Then

$$P(E_i) = P[4AC \leq B^2] = p, \text{ say},$$

and

$$\frac{\#\{i \leq N : E_i \text{ occurs}\}}{N} \to p$$

as $N \to \infty$ by the LLN.

Estimate: Estimate p by

$$\hat{p} = \frac{\#\{i \leq N : 4A_iC_i \leq B_i^2\}}{N}$$

with a large N.

Example: With $N = 10,000$,

$$\hat{p} = .628$$

Note: • Generate $30,000$ random numbers.

- Instantaneous.
- Error allowance .01 (Later).
- Exact: .627

Assessing The Error

Let

$$E_1, E_2, \cdots \text{ be independent,}$$
$$P(E_k) = p,$$
$$S = \#\{k \leq N : E_k \text{ occurs}\},$$

and

$$\hat{p} = \frac{1}{N}S = \frac{\#\{k \leq N : E_k \text{ occurs}\}}{N}.$$

Then

$$S \sim \text{Binomial}(N, p),$$
$$E(S) = Np,$$
$$D^2(S) = Npq,$$

and

$$S^* = \frac{S - Np}{\sqrt{Npq}} \approx \text{Normal}[0, 1].$$

Assessing Error-Continued

Let

$$\epsilon = \hat{p} - p$$

and

$$\epsilon^* = \frac{\hat{p} - p}{\sqrt{pq/N}}.$$

Then

$$\hat{p} - p = \sqrt{\frac{pq}{N}}\epsilon^*$$
$$\epsilon^* = S^* \approx \Phi.$$

So,

$$P[-2 \leq \epsilon^* \leq 2] \approx \Phi(2) - \Phi(-2) = .954$$

for large N. That is, with high probability

$$-2 \leq \epsilon^* = \frac{\hat{p} - p}{\sqrt{pq/N}} \leq 2,$$

or

$$|\hat{p} - p| = \sqrt{\frac{pq}{N}}|\epsilon^*| \leq 2\sqrt{\frac{pq}{N}}.$$

Assessing Error-Completed

Simple Bound: Use

$$pq = p(1 - p) \leq \frac{1}{4}.$$

So, with high probability,

$$|\hat{p} - p| \leq 2\sqrt{\frac{pq}{N}} \leq \frac{1}{\sqrt{N}}.$$

For example, if $N = 10,000$, then $1/\sqrt{N} = .01$.

Better Bound: Solve the inequality

$$|\hat{p} - p| \leq 2\sqrt{\frac{pq}{N}}.$$

Get

$$\ldots$$

Bad Idea (But widely recommended): Use

$$|p - \hat{p}| \leq 2\sqrt{\frac{\hat{p}\hat{q}}{N}}.$$

Remarks

Terminology: The interval

$$p \in \hat{p} \pm \frac{1}{\sqrt{N}}$$

is called a *(conservative, asymptotic) 95% confidence interval for p*.

Other Applications: Drug testings, opinion polls, etc.·.

Example: In a poll of $N = 1600$ voters, 55% favored candidate A over B. Allowance for error is $1/\sqrt{N} = .025$.

Example

Assessing Accuracy in the CLThm

Let

$$U_1, \cdots, U_{12} \sim^{ind} \text{Unif}(0, 1],$$
$$Z = U_1 + \cdots + U_{12} - 6,$$

and

$$F(z) = P[Z \le z]$$

then

$$F(z) \approx \Phi(z)$$

by the CLThm, since $\mu = 1/2$ and $\sigma^2 = 1/12$.

A Simulation Experiment: Generate

$$U_1, U_2. \cdots U_{12N}.$$

Let

$$Z_k = U_{12(k-1)+1} + \cdots + U_{12k} - 6$$

and

$$\hat{F}(z) = \frac{\#\{k \le N : Z_k \le z\}}{N}.$$

Results of the Simulation

$n = 12$

$N = 10,000$

z	$\hat{F}(z)$	$\Phi(z)$
-1.6	.058	.055
-1.2	.120	.115
-0.8	.214	.21.
-0.4	.346	.345
0.0	.496	.500
0.4	.654	.655
0.8	.789	.788
1.2	.884	.885
1.6	.944	.945

Simulating Other Distributions

Basic Fact: If

$$G \text{ is a cont. } \uparrow \text{DF},$$
$$U \sim \text{Unif}(0, 1],$$

then

$$Y = G^{-1}(U) \sim G,$$

since

$$\begin{aligned} P[Y \le y] &= P[G^{-1}(U) \le y] \\ &= P[U \le G(y)] \\ &= G(y). \end{aligned}$$

Note: • General Method

• Not always efficient.

Example

Exponential Distributions

If

$$G(y) = 1 - e^{-y},\ 0 \le y < \infty,$$
$$0 < u < 1,$$

then

$$G(y) = u$$

iff

$$e^{-y} = 1 - u,$$

so that

$$G^{-1}(u) = -\log(1 - u).$$

Thus

$$Y = -\log(1 - U) \sim G.$$

Note: Also, $-\log(U) \sim G$.